DOCUMENT RESUME

ED 216 259 CG 015 873

AUTHOR Brabeck, Mary Margaret
TITLE The Relationship between Critical Thinking Skills and Development of Reflective Judgment among Adolescent and Adult Women.
PUB DATE Aug 81
NOTE 33p.; Paper presented at the Annual Convention of the American Psychological Association (89th, Los Angeles, CA, August 24-26, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Adolescents; *Age Differences; *Cognitive Style; Comparative Analysis; *Critical Thinking; Females; Higher Education; *Intellectual Development; *Logical Thinking; Secondary Education; Young Adults
IDENTIFIERS *Reflective Thinking

ABSTRACT

Many educators believe that, regardless of innate abilities and acquired skills, students at different educational levels do not think in the same way. To account for these differences, 119 female students representing four educational levels (high school seniors to graduate students) completed the Watson-Glaser Critical Thinking Appraisal Form A and the Reflective Judgement Interview (RJI). The results supported previous reflective judgement studies in which RJI scores increased with educational level. This finding suggests that the development of reflective judgement is separate from and involves something other than the acquisition of thinking skills, although attainment of critical thinking skills is still necessary for the development of reflective judgement levels. (JAC)

The Relationship between Critical Thinking Skills and Development of Reflective Judgment Among Adolescent and Adult Women

Mary Margaret Brabeck

Boston College

A paper presented at The Annual Meeting of the American Psychological Association, Los Angeles California, 1981

It is almost a truism among experienced educators that regardless of innate abilities and acquired skills students at different educational levels do not think the same. There is probably some truth to the generalization that high school and first year college students want the instructor to think for them ("But what do you think? What is the correct belief?"); college juniors and seniors don't want anyone to think for them ("My answer is just as good as anyone else's. Everyone has a right to his own beliefs."); and graduate students want to think for themselves ("On the basis of the available evidence I believe the best answer . . .").

What accounts for these differences? Are they due to differences in acquired skill in logic and problem solving or is something else needed to account for these group differences? If students at different educational levels have attained the same thinking skills will they justify their beliefs in the same ways regardless of educational level? This study investigated these questions.

Intellectual development and the ability to think critically have long been stated aims of formal education. John Dewey (1916) took the importance of fostering good habits of thinking for granted: "All which the schools can or need do for pupils is to develop their ability to think" (Dewey, 1916, p. 1791). While there is little controversy about the value of this aim, definitions and descriptions of what constitutes good thinking vary.

In this paper two different theoretical descriptions of adolescent and young adult thinking are described and contrasted. The first, reflective judgment (See Appendix A; King, 1977;Kitchener, 1978; Kitchener and King, 1981) describes intellectual development as structural change occurring as the result of maturation and interaction with the environment. The second, critical thinking, (Follman and Miller, 1971; Glaser, 1941; Skinner, 1976) is described as a composite of skills involving logical reasoning and/or problem solving.

The constructs of reflective judgment and critical thinking share a number of commonalities. Aside from the obvious similarity in attention to intellectual or thinking ability they share an "attitude" toward thinking. Glaser characterizes one aspect of critical thinking as, "The persistent effort to examine any belief or supposed form of knowledge in light of the evidence that supports it" (Glaser, 1941, p. 6). In a similar way King (1977) describes high level reflective thinkers as being able to follow evidence through to a conclusion and as being "willing to both endorse the conclusion to which they have reasoned and to acknowledge that their views might be falsified by additional evidence obtained at a later time" (King, 1977, p. 16).

Both constructs also reflect an emphasis on the importance of knowledge of the principles of logic and skill in applying these principles. A number of writers have equated critical thinking with skill in logic (D'Angelo, 1970; Osborne and Reagan, 1973 and others). Others have included logic along with problem solving as components of critical thinking skills (Dolan,

1976; Skinner, 1976 and others). Watson and Glaser, (1964) authors of the Watson Glaser Critical Thinking Appraisal, (hereafter, WGCTA) operationally define high levels of critical thinking as, "The ability to correctly perform the universe of tasks represented by the five subtests" (p. 10). Research results on the WGCTA suggest that these specific "tasks" are primarily in the use of logic (Follman and Miller, 1971; Rust, Jones, and Kaiser, 1962).

Intellectual development, as described by the reflective judgment scheme, also includes logic abilities. "Logic," Elkind maintains, is "the bedrock upon which any epistomology must be built" (Piaget, 1967, p. 11). Subjects at lower levels of the reflective judgment scheme are depicted as using evidence illogically. That is, using evidence that contradicts the students' own beliefs. Subjects at the middle levels of reflective judgment disregard logical rules as they ignore them in reaching conclusions based on personal whim. At the highest level of reflective judgment logic provides important evidence. According to King (1977) judgments at level 7:

> may be differentiated from other types of judgments . . . by the way in which they are justified . . . justifications are explicitly based on reason, with reference to either empirical or logical evidence in support of the judgment . . . [A reflective] judgment is made by following evidence and logic through to a reasoned conclusion. (pp. 19-20)

In separate studies neither critical thinking nor reflective judgment levels have been demonstrated to be related to major in school (Simon and Ward, 1974; Skinner, 1971; Welfel, 1979). Verbal aptitude, general intelligence and educational level have been shown to be related to reflective judgment (King and Parker, 1978; Kitchener and King, 1981 ; Strange and King, in press; Welfel, 1978) as well as to critical thinking (Glaser, 1941; Little, 1972; Skinner, 1971; Westbrook and Sellers, 1967).

In summary, critical thinking and reflective judgment are constructs that characterize the thinking process in some similar ways. How is critical thinking different from reflective judgment? This is a critical question if reflective judgment is to stand as an independent construct. It is also a crucial issue for those who seek to understand the nature of thinking and intellectual development.

A major difference between the two constructs is seen in the different ways in which the acquisition or development of the construct is viewed. Though all people are assumed to be potentially capable of learning critical thinking skills, they can and must be taught. (Glaser, 1941, p. 19). Furst (1950) wrote, "Even though superior scholastic aptitude may enable students to accomplish various intellectual exercises without too much formal instruction in a given field, it is unlikely that students will reach the higher levels of performance in a subject field without the acquisition of certain specialized knowledge and instruction aimed at the cultivation of thinking skill" (p. 615). Glaser accepted Haslett's claim that reasoning of children

is basically the same as adults. This assumption allowed him to then claim that critical thinking skills can be taught at all ages (Glaser, 1941, p. 19).

Though a variety of approaches to teaching critical thinking have been tested and recommended, (Decaroli, 1973; Kemp, 1963; Shatin and Opdyke, 1967, among others), the assumption about critical thinking is the same in all these studies: critical thinking is a skill that can be taught. This is in contrast with the developmental approach of reflective judgment that views cognitive development as structural change occurring through interaction of maturation and experience. While reflective judgment development involves increasing skill in logic as well as use of empirical evidence to justify beliefs, the theory claims also to meet the criterion of a stage theory. Changes in reflective judgment are not viewed as accumulation of new skills, nor as refinements of previously learned skills, but as reorganizations or structural transformations that result in qualitatively different epistemological and metaphysical assumptions. Thus, persons reasoning at any given level organize their experience in ways that are qualitatively different from and not reducible to lower stages. Further, these qualitative changes occur in an age-related pattern of stages or levels.

It was hypothesized in this study that critical thinking skills would be included in and limit but would not fully account for reflective judgment development. Low critical thinking subjects were predicted to be homogeneously low in RJI scores, while high critical thinking subjects were predicted to be higher

and more heterogeneous in their reflective judgment levels. In other words, failure to acquire skills of critical thinking was predicted to limit reflective judgment level. Greater variability of reflective judgment scores among high scoring critical thinkers, (some scoring high, some low on reflective judgment), would support the claim that critical thinking skills alone are not sufficient for attainment of high reasoning levels.

In an attempt to provide a stricter test of the hypothesized relationship between critical thinking skills and reflective judgment development, subjects were selected to represent extremes of high and low critical thinking abilities. High and low groups were defined by absolute values obtained from WGCTA norm information, rather than from a score determined as high or low relative to the sample.

Method

Subjects

The subjects for this study were 119 female students enrolled in Catholic institutions at four educational levels. All students except one Black woman were Caucasion. High school seniors were all 18 or 19 years old; college sophomores ranged in age from 19 to 35 years old; college seniors ranged from 21 to 37 years old; graduate students ranged from 22 to 47 years old. To identify subjects for this study an initial pool of 392 volunteer female students were tested on the WGCTA Form A.

Two groups (high and low critical thinkers) were identified as follows: Norms accompanying the WGCTA Forms A and B were used to determine the expected high and low 15% for this sample.

These scores were respectively 60 and 52. The fifteen highest scoring high school seniors who met the criterion of 60 or above were selected and the other three groups matched within 1 point to tneir scores. A similar process was used to identify the low group. However, only 14 graduates scored at 52 or below. These students were selected and the other three groups matched to their scores.

Measures

Three instruments were administered to each subject in the following order: A Biographical Data Sheet, The Watson-Glaser Critical Thinking Appraisal Form A, and The Reflective Judgment Interview.

Biographical Data Sheet: The biographical data sheet was used to obtain demographic information. The instrument consists of three pages and takes about five minutes to complete. Data from this questionnaire were obtained in order to describe the sample.

The Watson-Glaser Critical Thinking Appraisal, Form A: Of the tests that have been developed to measure critical thinking, only a few have established norms for college/adult populations. One of the most widely used of these instruments is the Watson-Glaser Critical Thinking Appraisal.

The WGCTA, Form A is a refinement of earlier forms. Subjects are timed at forty minutes and are instructed in the standardized directions to attempt each question, to move rapidly through the test, and to use the remaining time to go back over any part of the test. The test includes five subtests, 16 items

in each: Inference, Recognition of Assumptions, Deduction, Interpretation, and Evaluation or Arguments. "Right" answers were originally determined by submitting each item to a jury of 15 people trained in logic and scientific method.

Reflective Judgment Interview: All four dilemmas were presented by one of two trained interviewers to each subject individually. Dilemmas were presented in random order to reduce the effects of fatigue and ordering. The tapes of the 119 interviews were transcribed into four separate typed protocols, one for each dilemma for a total of 476 protocols. Blind rating of the interviews was assured by editing from the taped interviews any reference to age or year in school, identifying each of the 476 protocols with a randomly assigned number and by random presentation of the transcribed interview to the raters. Each of the protocols was scored separately by two trained and certified raters, neither of whom were involved in interviewing, transcribing, or editing the interviews.

In accord with procedures used in all previous studies of reflective judgment, each protocol was assigned three numbers indicating the reflective judgment level or levels most represented in the protocol. A subject's reflective judgment score is the average of all scores assigned by two raters.

Procedure

All 119 subjects participating in the interview were first asked to read and sign a statement of informed consent. They were then given the semi-structured RJI which was tape recorded. After the interview, subjects were asked to rate the dilemmas in

order of their familiarity with the topic and degree of importance they assigned to each issue. Administration of the interviews took three months during the winter of 1980.

RESULTS

Reliability of the Measures

The reliability of the RJI over all dilemmas for two judges' independent ratings (Pearson product-moment correlation) was .77. Rater reliabilities obtained for each dilemma ranged from .53 to .62; all rater reliabilities were highly significant ($p < .001$). Inter-rater agreement, the percent of times the judges made the same judgments about each protocol and were discrepant by less than a stage was .76. Correcting for change agreement the Lawlis and Lu (1972) t coefficient was .76. Agreement levels within dilemmas ranged from .73 to .78.

Internal consistency of the RJI was evaluated by several means. Pearson product-moment correlations between all pairs of dilemmas were between .35 and .47, $p < .001$. Dilemma-total correlations, comparing subjects' scores on one dilemma with the other three dilemmas were between .52 and .59, $p < .0001$. Cronbach's alpha co-efficient was computed to measure the internal consistency of the RJI measure; the overall alpha level was .75. Alpha levels increased with educational level; the high school group obtained an alpha of .64, college sophomores yielded .68, college seniors .72, and graduates .70.

No differences were found in RJI scores of subjects interviewed by the two interviewers. No relationship was found between RJI scores and subjects' reported familiarity with the

dilemmas nor with the importance attributed to the different dilemmas.

WGCTA subtest intercorrelations (Pearson product-moment coefficients) ranged from .29 to .50, $p. < .001$. Kuder-Richardson - 18 reliabilities for each subtest ranged from .38 to .79 when computed by educational level ($N = 30$). Considering the few number of items ($N = 16$) in each subtest, these reliabilities are respectable. Composite KR-18 reliability coefficients when computed by educational level ranged from .81 (graduate group) to .89 (high school group). Composite scores are coefficient alphas based on subtest variability. Item-total correlation coefficients corrected for effects of a given item ranged from .46 to .58, $p. < .001$. The overall internal consistency of the WGCTA, the alpha level, was .76. These data suggest the instruments used in this study were reliably assessing critical thinking and reflective judgment.

Distribution of Reflective Judgment Scores

The distribution of mean Reflective Judgment Interview scores for each educational level are presented in Table 1.

Insert Table 1 about here

The scores range from 2.5 to 5.3 with high school seniors and college sophomores scoring lower ($\bar{X} = 3.4$) than college seniors ($\bar{X} = 3.7$), and college seniors scoring lower then graduate students ($\bar{X} = 4.0$). An analysis of variance was conducted to test for differences among educational levels. A

significant main effect was found for educational levels, F (3, 116) = 13.31 $p < .0001$.

The Waller-Duncan K-ratio was used to follow up on the analysis of variance and examine differences in mean RJI scores among each educational level. Graduate students and college seniors were found to differ significantly $F(3, 112) = 13.31$, $p < .05$, from each other and from college sophomores and high school students. No difference was found between high school seniors and college sophomores. The mean RJI level of high school seniors in this sample, 3.4, is over half a stage higher than that of the only other high school group measured for reflective judgment. Kitchener and King (in press) reported a mean of 2.8 for their sample of high school juniors from public institutions. It may be that the type of school subjects attend affects the development of RJI level. This would be consistent with Astin's (1977) observations about the effect of the size of an institution on students' cognitive and affective development. He reported that smaller, private colleges with greater opportunities for involvement in campus activities and programs and more faculty student interaction had a greater impact on both cognitive and affective dimensions. However, the college sample of this study, also from small, private institutions did not show a similar increase in RJI levels.

Comparison of the means of Scholastic Aptitude Test scores shows that high school students scored higher ($\bar{X}$ = 453.3) than college sophomores ($\bar{X}$ = 433.3) on SAT Verbal tests. Verbal SAT test have a moderate and significant correlation with the RJI

overall score (r = .53 $p < .01$). This finding raises speculation about whether the high school students scored higher on RJI than high school students of previous studies because of higher SAT scores. It may also explain why no difference between high school seniors and college sophomores on RJI was found, i.e., the increased verbal aptitude among high school seniors may have made them more equal to college sophomores in RJI. However, these speculations must be interpreted in light of the fact that previous studies (Kitchener, 1977; Welfel, 1979) which reported differences among educational level groups in verbal ability as measured by the Concept Mastery Test found those differences remained when CMT was covaried out. Further, Welfel (1978) found no significant correlations between RJI and PSAT and Mines (1980) found differences in RJI scores among educational levels remained when ACT scores were covaried out. When SAT scores from high school and college subjects in this study were covaried out, obtained differences in reflective judgment scores among educational levels remain $F(2,74) = 17.30$, $p < .0001$. The RJI scores of college seniors remained significantly different from college sophomores and high school seniors. These findings support previous claims that reflective judgment is separate from verbal ability.

The variability among RJI scores increases from high school students ($S.D.$ = .40) to graduate students ($S.D.$ = .57). Cochran's C was used to test for the difference in variability of students among educational levels. The graduate students with the highest variance were found to differ significantly from the

other groups ($p < .05$). Since the internal consistency of the RJI (alpha levels) was shown to increase with educational level, differences in the variability of subjects at different educational levels cannot be attributed to the unreliability of the RJI for the graduate group. These data suggest that as students mature educationally, they vary more in terms of their reflective judgment levels.

The Relationship between the Measures

The overall correlation between Reflective Judgment Interview scores and critical thinking scores was .40, $p < .001$. These coefficients ranged from .37 to . 48 when computed by educational level and remained significant, $p < .05$. Table 2 presents the obtained correlations between RJI scores and WGCTA scores.

Insert Table 2 about here

The proportion of variance shared by WGCTA scores and RJI scores is .16. Given the good internal consistency of the two instruments (WGTA = .76; RJI = .75) the obtained correlation between the two is lower than one would expect if the two instruments were measuring the same construct. These data support the claim that reflective judgment and critical thinking as measured in this study are separate constructs.

Differences in Means between High and Low Critical Thinking Subjects

Overall mean RJI scores were higher for high critical thinking subjects than for low critical thinking subjects at each educational level. These data are presented in Table 3.

Insert Table 3 about here

An analysis of variance was conducted to test for the differences between high and low critical thinking subjects on RJI scores, (F (1, 118) = 30.49, $p < .001$). The interaction between educational level and high versus low critical thinking was not significant. This indicates that no single educational level group completely accounts for the obtained difference between high and low critical thinkers.

To follow up on these results and examine the differences between high and low critical thinking subjects' mean scores of reflective judgment at each educational level, Dunn's multiple comparison procedure was used. Only the college sophomore sample was not significantly different at the .05 level. Overall mean differences and those obtained for each educational level were in the hypothesized direction with high critical thinking subjects scoring higher on the RJI than low critical thinking subjects.

Of the 59 subjects in the low critical thinking group of this study only 1 high school senior and 2 graduate students were rated at a RJI level above 4.0. No low scoring subject was rated higher than 4.2.

To examine if there were any evidence of higher RJI levels of reasoning, the individual dilemma scores of this study were

studied. Out of 476 protocols, 32 were assigned scores of 4.5 or above. Only one dilemma of a low scoring critical thinking subject was rated at 4.5; no low critical thinking subjects obtained a score above 4.5 on any individual dilemma. Given the small number of RJI ratings above 4.5 obtained in this study, these results are not conclusive. However, they do give preliminary support to the claim that attainment of critical thinking skills precedes and may limit development of reflective judgment levels for the middle levels of the scheme.

Differences in Variability between High and Low Critical Thinking Subjects

Even though the range of the high critical thinking group was more restricted in WGCTA scores (range = 60 - 71) than the low critical thinking group (range = 32-52) the equality of variance test indicated that there was significantly greater variability among the RJI scores of the high critical thinking group, F (58, 59) = 1.69, $p < .05$. Overall low critical thinking subjects yielded a variance on RJI of .16 and overall high critical thinking subjects yielded a variance on RJI scores of .27. When the differences in variance at each educational level were tested, differences among high and low critical thinking subjects were found to be significant only in the college sophomore group F (14,14) = 3.42, $p < .05$). However, the lack of significance at the other educational levels must be interpreted in light of the small number of subjects.

These data suggest further analyses with larger cell N's is necessary before making any definate statements about the

differences in variance at the educational levels. However, overall in this study people tend to vary more in their reflective judgment levels with attainment of higher critical thinking skills as hypothesized.

One possible explanation for the obtained difference in variability between high and low critical thinking groups might be attributable to differences in the reliability of the RJI measure. That is, differences in variability would be expected if the RJI was less reliable for the graduate sample. Cronbach's alpha coefficient was computed as a measure of RJI internal consistency for the high and the low groups. Internal consistency was found to be higher for the high critical thinking group (.72) than the low group (.66). The greater variability of RJI scores among high critical thinking subjects cannot be attributed to differences in the reliability of the instrument.

Discussion

The results of this study support previous reflective judgment studies (Brabeck, et al) that indicate that RJI scores increase with educational level. Subjects for this study were matched across educational levels on critical thinking scores, so differences in reflective judgment levels cannot be attributed to differences in critical thinking skills as measured by the Watson Glaser Critical Thinking Appraisal. This finding suggests that development of reflective judgment is separate from and involves something other than the acquisition of thinking skills. Even though these skills were equal across educational levels, younger less educationally advanced students were more likely to look to

authorities for certain answers that were assumed to exist. More educationally advanced students were more likely to consider a number of explanations feasible and justifiable and were more likely to view knowledge as uncertain and contextually relative.

This is not to suggest, however, that reflective judgment is unaffected by acquisition of critical thinking skills. The results of this study suggest, rather, that attainment of critical thinking skills is necessary for development of reflective judgment levels. High critical thinkers out-scored low critical thinkers and yielded greater variability on the RJI. No low critical thinkers scored above RJI level 4.2.

It appears, however, that even though students may master critical thinking skills, such skills do not ensure development of higher levels of reflective judgment. In this study subjects who were able to identify appropriate deductions and inferences, to recognize assumptions and to evaluate arguments, did not all score at higher RJI levels. Even masters level graduate students who scored in the top 15% of the WGCTA Form A adult norms had a RJI mean of only 4.2 and ranged from 2.8 to 5.3. Graduate students high in critical thinking skills did not differentiate between good and bad evidence, could not form a reasoned synthesis from conflicting data, and frequently used whim rather than evidence to justify their beliefs. The RJI appears to be measuring processes of reasoning that are very different from and perhaps more than the thinking skills measured by the WGCTA.

These findings have significance for those engaged in higher education when the results are viewed from the perspective of the

philosophical description of the aims of education as proposed by R.S. Peters (1966). Peters argues that an aim involves a process rather than a product. The word "aim" connotes that a person may fall short of achievement of something because of the difficulty involved in the task. Formulation of the aims of education is important because such a statement adds structure and coherence to activities associated with reaching the aim (Peters, 1966, pp. 5-7).

For Peters, the aim of education is broader than acquisition of skills: "For a man to be educated it is insufficient that he should possess a mere know-how or knack. He must have also some kind of conceptual scheme to raise this above the level of a collection of disjointed facts: (Peters, 1966, p. 8, emphasis added). In reaching this aim, Peters says, education "picks out no particular activity or process. Rather, it lays down criteria to which activities or processes must conform" (Peters, 1966, p. 3). Herein lies an important difference between reflective judgment and critical thinking as defined and investigated in this study.

The structural changes, the development of the epistemological and metaphysical assumptions about knowledge, which lead one to form an attitude toward the nature of truth, knowledge, evidence and authorities - these are the defining characteristics of the reflective judgment scheme. It may be that the contribution that reflective judgment theory makes is its description of these developing "conceptual schemes" that raise our knowledge "above the level of disjointed facts."

Statement of the aim of the development of thinking as reflective judgment then serves to unify activities such as teaching specific skills (deductive reasoning, evaluation of arguments, logic, etc.), assigning specific tasks, and requiring mastery of specified levels of proficiency that are appropriate to an individual's level of developing reflective judgment. The composite of skills in deduction, inference, recognition of assumptions, interpretation and evaluation, (the subtests of the WGCTA) found in this study to be associated with advanced levels of reflective judgment, can be viewed as "activities or processes" that are necessary but not sufficient for attainment of the highest levels of reflective judgment.

The results obtained from this sample were expected to be lower than those obtained from samples that included male and Ph.D. level students. Some studies have reported that males are significantly higher on RJI than females although the magnitude of the differences was not great (Lawson, 1980; Mines, 1980; Strange, in press). More educationally advanced Ph.D. students might be expected to score higher than masters level students. However, except for the results of King (1977) and Kitchener's (1978) original studies, the results of RJI scores for graduate samples reveal a disturbing lack of subjects scoring in the highest levels of reflective judgment (Brabeck et al., 1981).

It may be that subjects are capable of reasoning at levels 6 and 7 but this reasoning is not accurately reflected in the RJI score given the subject. Alternatively it can be argued that the highest scores obtained from response to any given protocol

define the upper limits of a person's reasoning abilities and are, therefore, a more accurate statement of the level of thinking that the person is capable of producing. To test this idea in the future, RJI score could be computed using only the two dilemmas on which each individual scored highest; the average of the four ratings (2 dilemmas X 2 raters) could be used as the RJI score.

A final explanation that warrants investigation is that the type of reasoning characterized as reflective judgment level 7 is not common in today's world, including the world of today's university. It may be that Horace's (in Kraemer, 1936) challenge to "seek the truth in the groves of Academe" (p. 380, line 45) is no longer appropriate. It may be that the universities reflect a larger pattern in society. This possibility makes it even more important that educators strive to understand the nature of the "conceptual scheme" that as Peters (1966, p. 8) says, raises knowledge and skill "above the level of a collection of disjointed facts". Reflective judgment may prove useful in describing the appropriate educational aim of the development of this conceptual scheme. Teaching critical thinking skills may also prove to be one of the appropriate activities to promote achievement of that aim.

References

Astin, A. W. Four Critical Years. San Francisco: Jossey Bass Inc., 1977.

Brabeck, M., King, P., Kitchener, K., Mines, R., Welfel, E. Learning to make judgments: How students defend what they believe. A presentation made at the annual meeting of the American College Personnel Association, Cincinatti, 1981.

D'Angelo, E. Philosophers as critical thinking consultants. School and Society, 1970, 98, 166.

DeCaroli, J. Critical thinking. Social Education, 1973, 37, 67-69.

Dewey, J. Democracy and Education. New York: The MacMillan Company, 1916.

Dolan, S. The effect of value-conflict discussion on the critical thinking and reading of adults. (Doctoral dissertation, University of Pittsburgh, 1976). Dissertation Abstracts International, 1976, 37, 2737A.

Follman, J. and Miller, W. Statistical analysis of three critical thinking tests. Educational and Psychological Measurement, 1971, 31, 519-520.

Furst, E. J. The relationship between tests of intelligence and tests of critical thinking and knowledge. Journal of Educational Research, 1950, 43, 614-625.

Glaser, E. M. An experiment in the development of critical thinking, contribution to Education #843. New York: Bureau of Publications Teachers College, Columbia University. 1941.

King, P. M. The development of reflective judgment and formal operational thinking in adolescents and young adults. (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 1977, 38, 7233A.

King, P. William Perry's theory of intellectual and ethical development in Knefelkamp, L., Widick, C., and Parker, C., (Eds.) Applying new developmental findings. San Francisco: Jossey-Bass, 1978.

King, P., and Parker, C. Assessing intellectual development in the college years: A report on the instructional improvement project, 1976-1977. Unpublished improvement project, 1976-1977. Unpublished manuscript, University of Minnesota, 1978.

Kitchener, K. S. Intellectual development in late adolescents and young adults: Reflective judgment and verbal reasoning. (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 1978, 39, 936B.

Kitchener, K., and King, P. Reflective judgment: Concepts of justification and their relationship to age and education. Journal of Applied Developmental Psychology, 1981, 2, 89-116.

Kitchener, K., and Kitchener, R. The development of natural rationality: Can formal operations account for it? Human Development, in press.

Kraemer, C. J. (Ed), Complete Works of Horace, New York: Random House, 1936.

Lawson, J. The relationship between graduate education and the development of reflective judgment: A function of age or

educational experience. (Doctoral dissertation, University of Minnesota, 1980). Dissertation Abstracts International, 1981, 41, 4655A.

Little, T. L. The relationship of critical thinking ability to intelligence, personality factors, and academic achievement. (Doctoral dissertation, Memphis State University, 1972). Dissertation Abstracts International, 1970, 31, 680A.

Mines, R. Levels of intellectual development and associated critical thinking skills in young adults. (Doctoral dissertation, University of Iowa, 1980. Dissertation Abstracts International, 1981, 41, 1495A.

Ness, J. H. The effects of a beginning speech course on critical thinking ability. Doctoral dissertation, University of Minnesota, 1967). Dissertation Abstracts International, 1969, 28, 5171-5162.

Osborne, A. R. and Reagon, G. M. Logical reasoning: An educational goal. Theory into Practice, 1973, 12, 263-265.

Peters, R. S. Ethics and education. Chicago: Scott Foresman and Company, 1966.

Piaget, J. Six psychological studies, New York: Vintage Books, 1967.

Rest, J. R. New approaches in the assessment of moral judgment. In T. Lickona (Ed.) Moral Development and Behavior. New York: Holt, Rinehart and Winston, 1976.

Rust, V. I., Jones, R. S., and Kaiser, H. F. A factor analytic study of critical thinking. The Journal of Educational Research, 1962, 55, 253-259.

Shatin, L. and Opdyke, D. A critical thinking appraisal and its correlate. Journal of Education, 1967, 42, 787-792.

Skinner, S. B. Cognitive development: A prerequisite for critical thinking. The Clearing House, 1976, 49, 292-299.

Skinner, S. B. The myth of teaching for critical thinking. The Clearing House, 1971, 45, 373-376.

Strange, C., and King, P. Intellectual development and its relationship to maturation during the college years. Journal of applied developmental psychology, in press.

Watson, G. and Glaser, E. Watson-Glaser Critical Thinking Appraisal Manual. New York: Harcourt, Brace, and World, Inc., 1964.

Welfel, E. R. The development of reflective judgment: Its relationship to year in college, major field, academic performance and satisfaction with major among college students. Unpublished doctoral dissertation, University of Minnesota, 1979.

Westbrook, B. W. and Sellers, J. R. Critical thinking, intelligence, and vocabulary, Educational and Psychological Measurement, 1967, 27, 443-446.

Table 1

Distribution of Mean Reflective Judgment Interview Scores

Over All Dilemmas for Each Educational Level

(by percent*)

Mean Reflective Judgment Scores	High School Seniors	College Sophomores	College Seniors	Masters Level Graduates
2.0				
2.5	17** (4)***	17 (4)		3 (1)
3.0	40 (10)	57 (14)	30 (8)	17 (4)
3.5	33 (8)	23 (6)	50 (13)	41 (10)
4.0	10 (3)		17 (4)	26 (6)
4.5		3 (1)	3 (1)	10 (3)
5.0				3 (1)
5.5				
6.0				
	N = 30	N = 30	N = 30	N = 29
	$\bar{X}$ = 3.4	$\bar{X}$ = 3.4	$\bar{X}$ = 3.7	$\bar{X}$ = 4.0
	SD = .40	SD = .42	SD = .42	SD = .57
	Range = 2.5-4.1	Range = 2.8-4.9	Range = 3.1-5.0	Range = 2.8-5.3

* Numbers may not sum to 100 because of rounding

** Percent of the particular educational level

*** Percent of total sample (N = 119)

Table 2

Correlation Coefficients Between RJI Dilemmas and

Overall Scores and WGCTA Subtests and Composite Scores

(N = 119)

	Reflective Judgment Interview Dilemmas				
WGCTA Subtests	Egyptian Pyramids	News Reporting	Chemicals in Food	Creation/ Evolution	Overall RJI
WG1	.15	.17	.27	.16	.24
	$p < .05$	$p < .033$	$p < .002$	$p < .041$	$p < .004$
WG2	.35	.26	.36	.28	.38
	$p < .003$	$p < .002$	$p < .001$	$p < .001$	$p < .001$
WG3	.13	.24	.27	.27	.30
	$p < .078$	$p < .004$	$p < .002$	$p < .001$	$p < .001$
WG4	.29	.28	.15	.30	.34
	$p < .001$	$p < .001$	$p < .054$	$p < .001$	$p < .001$
WG5	.03	.17	.18	.13	.17
	$p < .369$	$p < .034$	$p < .024$	$p < .077$	$p < .033$
WGComp.	.23	.31	.34	.31	.40
	$p < .01$	$p < .0006$	$p < .0001$	$p < .0005$	$p < .001$

Table 3

Overall Means and Standard Deviations of Reflective Judgment Interview Scores for Low and High Critical Thinking Groups by Educational Level

Educational Level and Critical Thinking Group	Mean	SD	Range
High School Low CT	3.2	.41	2.5 - 4.1
High School High CT	3.6	.33	2.9 - 4.1
College Sophomore Low CT	3.2	.27	2.8 - 3.8
College Sophomore High CT	3.5	.50	3.0 - 4.9
College Senior Low CT	3.5	.34	3.1 - 4.0
College Senior High CT	4.0	.38	3.5 - 5.0
Master's Level Low CT	3.7	.42	2.8 - 4.2
Master's Level High CT	4.2	.57	3.4 - 5.3

Appendix A

The Development of Reflective Judgment (Kitchener & King, 1981)

Stage	A) Metaphysical Assumptions	B) Epistemological Assumptions	Concepts of Justification
1	There is an objective reality which exists as the individual sees it. Reality and knowledge about reality are identical and known absolutely through the individual's perceptions.	Knowledge exists absolutely. One's own views and those of authorities are assumed to correspond to each other and to absolute knowledge. Knowledge is gained through the individual's perceptions and prior teaching.	Beliefs simply exist; they are not derived and need not be explained. Differences in opinion are not perceived, and justification is therefore unnecessary.
2	There is an objective reality which is knowable and known by someone.	Absolute knowledge exists, but it may not be immediately available to the individual. It is, however, available to legitimate authorities.	Beliefs either exist or are based on the absolute knowledge of a legitimate authority.
3	There is an objective reality, but it cannot always be immediately known, even to legitimate authorities. It is possible to attain knowledge about this reality, but our full knowledge of it is as yet incomplete and therefore uncertain.	Absolute knowledge exists in some areas, but in others it is uncertain, at least temporarily. Even authorities may not have certain knowledge, and therefore cannot always be depended upon as sources of knowledge. Knowledge is manifest in evidence which is understood in a concrete, quantitative way such that a large accumulation of evidence will lead to absolute truth.	Beliefs either exist or are based on an accumulation of evidence that leads to absolute knowledge. When such evidence is not available, individuals claim that while waiting for absolute knowledge to become available, people can temporarily believe whatever they choose to believe.
4	There is an objective reality, but it can never be known without uncertainty. Neither authorities, time or money nor a quantity of evidence can be relied upon to ultimately lead to absolute knowledge.	Absolute knowledge is for practical reasons impossible to attain, and is therefore always uncertain. There are many possible answers to every question but without certainty and a way to adjudicate between answers, there is no way to decide which one is correct, or even whether one is better than another. Knowledge is idiosyncratic to the individual.	Beliefs are justified with idiosyncratic knowledge claims and on idiosyncratic evaluations of data ("What is true for me, but not necessarily for anyone else"). The individual is the ultimate source and judge of his or her own truth.

Appendix A (cont.)

Stage	A) Metaphysical Assumptions	B) Epistemological Assumptions	Concepts of Justification
5	An objective understanding of reality is not possible since objective knowledge does not exist. Reality exists only subjectively and what is known of reality reflects a strictly personal knowledge. Since objective reality does not exist, an objective understanding of reality is not possible.	Knowledge is subjective. Knowledge claims are limited to subjective interpretations from a particular perspective based on the rules of inquiry and of evaluation compatible with that perspective.	Beliefs are justified with appropriate decision rules for a particular perspective or context, e.g., that a simpler scientific theory is better than a complex one.
6	An objective understanding of reality is not possible since our knowledge of reality is subject to our own perceptions and interpretations. However, some judgments about reality may be evaluated as more rational or based on stronger evidence than other judgments.	Objective knowledge is not possible to attain because our knowledge is based on subjective perceptions and interpretations. Knowledge claims can be constructed through generalized principles of inquiry and by abstracting common elements across different perspectives. The knowee must play an active role in the construction of such claims.	Beliefs are justified for a particular issue by using generalized rules of evidence and inquiry. However, since our understanding of reality is subjective, any such justification is limited to a particular case, time or issue.
7	There is an objective reality against which ideas and assumptions must ultimately be tested. Despite the fact that our knowledge of reality is subject to our own perceptions and interpretations, it is nevertheless possible, through the process of critical inquiry and evaluation, to determine that some judgments about that reality are more correct than other judgments.	Objective knowledge is possible to attain. Knowledge is the outcome of the process of reasonable inquiry. The process of inquiry, however, may not always lead to correct claims about the nature of reality since the process itself is fallible. Knowledge statements must be evaluated as more or less likely approximations to reality and must be open to the scrutiny and criticisms of other rational people.	Beliefs reflect solutions that can be justified as most reasonable using general rules of inquiry or evaluation. Criteria for evaluation may vary from domain to domain (e.g., religion, literature, science), but the assumption that ideas, beliefs, etc. may be judged as better or worse approximations to reality remains constant.

33